Preprints of the 21st IFAC World Congress (Virtual)
Berlin, Germany, July 12-17, 2020

The Evolution of Cooperation and Reward in a Corrupt Environment

Linjie Liu and Xiaojie Chen *

* *School of Mathematical Sciences, University of Electronic Science and Technology of China, Chengdu 611731, China (e-mail: xiaojiechen@uestc.edu.cn).*

Abstract: The maintenance of cooperative behaviors in many complex social systems has always been a great challenge. It has been suggested that costly punishment and reward can both facilitate the evolution of cooperation. Recent theoretical work, however, reveals that the positive role of centralized punishment in promoting cooperation has been challenged when violators can bribe centralized authorities to escape from sanctions. Naturally, the question arises as to how cooperation evolves when defectors can bribe rewarders for getting the reward. Here, we propose an evolutionary game theoretical model in which defectors can choose to bribe the rewarders probabilistically, and meanwhile rewarders will stochastically receive bribes from defectors in the public goods game. We theoretically study deterministic dynamics in infinite populations, and find that cooperators, defectors, and rewarders can coexist and the fraction of each strategist in the population remains unchanged in the coexistence state. Furthermore, we numerically investigate stochastic dynamics in finite populations, and reveal that cooperative behaviors can be maintained since the population can spend most of the time in the region where cooperators, defectors, and rewarders coexist.

Keywords: Evolutionary game theory, public goods games, corruption, reward, cooperation, stochastic dynamics

1. INTRODUCTION

The emergence of cooperative behaviors in the multi-agents system is an attractive research topic (see Nowak et al. (2004); Nowak (2006); Perc et al. (2017); Govaert et al. (2017); Kawano et al. (2017); Xiao et al. (2020)). A simple definition of cooperation is that one individual pays a cost for another to receive a benefit, which renders that cooperation is an irrational choice and will be eliminated by natural selection. However, cooperation is a ubiquitous phenomenon in nature and human society. In order to solve this inconsistency, several incentive mechanisms have been proposed in recent years for explaining the emergence of public cooperation, such as reward and costly punishment (see Rand et al. (2009); Gao et al. (2015); Chen et al. (2015); Liu et al. (2019); Wang et al. (2019)).

Many experimental and theoretical studies have confirmed that individuals are willing to incur costs to punish uncooperative individuals (see Fehr and Gächter (2002); De Quervain et al. (2004); Chen et al. (2014); Wang et al. (2018)). However, the effect of centralized punishment in promoting cooperation has been challenged by recent theoretical research which shows that the existence of corruption under which free-riders bribe corrupt officials to avoid punishment has undermined the positive role of pool punishment in cooperation (see Abdallah et al. (2014)). In reality, this form of corruption is widespread, and possibly causes a large decrease in public good provision (see Huang et al. (2018)). Recent behavioral experiments have further confirmed that empowering leaders reduce cooperative contributions once they fall into corruption, which is contrary to findings about typical institutional punishment (see Muthukrishna et al. (2017)).

Reward is an established alternative to punishment, although there has been relatively little research in the past. Different from costly punishment, reward, as a positive incentive, is used to increase contributors' welfare (see Sigmund et al. (1993); Szolnoki and Perc (2010); Sasaki and Unemi (2011); Szolnoki and Perc (2012)). Previous literatures assume that punishment mechanism is vulnerable to corruption because its purpose is to reduce the payoff of free-riders (see Abdallah et al. (2014); Huang et al. (2018); Liu et al. (2019)). However, there is little literature to investigate whether the effectiveness of pool reward in promoting cooperation can be undermined when defectors have the opportunity to enhance their benefit by bribing rewarders.

With this in mind, we thus introduce corrupt rewarders and defectors into the public goods game (PGG) with pool reward. We assume that defectors bribe rewarders probabilistically in order to get the bonus, and meanwhile pool rewarders receive bribes from defectors stochastically. We theoretically analyze replicator dynamics in infinite populations, and reveal that cooperators, defectors, and rewarders can coexist steadily in the population. Furthermore, we provide numerical examples to verify our theoretical results. Finally, we numerically investigate stochastic dynamics in finite populations with behavioral mutations.

⋆ This research was supported by the National Natural Science Foundation of China (Grants Nos. 61976048 and 61503062)

Since stochasticity is present, there is no fixation and the population can spend most of the time in the coexistence states of cooperators, defectors, and rewarders. Our results clearly suggest that cooperation can indeed be maintained no matter whether the population size is finite or infinite.

2. MODEL

2.1 PGG

We use the PGG to investigate the evolution of cooperation. From time to time, N players are randomly drawn from the population to form a group and play a one-shot PGG. Each player chooses whether to participate in the PGG as a cooperator (C) or defector (D). Each C pays a cost c to contribute to the common pool, whereas D refuses to do so. The accumulated contribution is multiplied by an enhancement r (with $1 < r < N$), and shared evenly among all participating players of the group.

2.2 Pure pool reward

In the pool rewarding stage, each player can decide whether or not to contribute an amount G_R to the reward pool before participating in the PGG, and the contributors are responsible for rewarding players who contribute to the PGG. Considering that in real society, it could be inefficient or even impossible to perform both cooperation and reward simultaneously due to the limitation of resource or time, here we assume that players in the group can only play one of three roles, pure cooperator (C, performs cooperating but never rewarding), pure defector (D, performs neither cooperating nor rewarding), or pure rewarder (R, performs rewarding but never cooperating). Then each C player can receive α from each R player of the group. It's worth mentioning that some defectors choose to bribe rewarders probabilistically in order to enhance their own earnings, and meanwhile some rewarders choose to accept bribes stochastically in real society. In this way, we assume that defectors will choose to pay bribes h ($h < \alpha$) to each corrupt rewarder with probability p. With probability $1 - p$, defectors do not offer bribes to each corrupt rewarder. Similarity, rewarders decide to accept bribes b from each corrupt defector with probability q, while with $1 - q$ probability, rewarders are unwilling to accept bribes. The concrete parameters are shown in Table 1.

Table 1. Notation symbols and meanings

Symbols	Meanings
N	Group size
r	Multiplication factor
c	Cost of cooperation
G_R	Cost of pool reward
α	Bonus from pool rewarders
h	Bribes provided by the defectors
b	Bribes received by the rewards
p	Probability that defectors fall into corruption
q	Probability that rewarders fall into corruption

Based on these assumptions, the payoffs of a pure cooperator, C, a pure defector, D, and a pure rewarder, R, who play the PGG with N_C pure cooperators, N_R pure rewarders, and N_D pure defectors ($N_C+N_R+N_D = N-1$) in one group can be, respectively, given by

$$\pi_C = \frac{rc(N_C+1)}{N} - c + N_R\alpha, \tag{1}$$

$$\pi_D = \frac{rcN_C}{N} - pqN_Rh + pqN_R\alpha, \tag{2}$$

$$\pi_R = \frac{rcN_C}{N} - G_R + pqN_Db, \tag{3}$$

where $N_R\alpha$ denotes the expected reward from rewarders of the group.

We assume that individuals tend to copy others whenever these appear to be more successful. Then the social learning approach is suitable for investigating strategy evolution as time goes by. In the following sections we will respectively study evolutionary dynamics in infinite and finite well-mixed populations.

3. EVOLUTIONARY DYNAMICS IN INFINITE WELL-MIXED POPULATIONS

For studying the evolutionary dynamics of different strategies in infinite well-mixed populations, we choose to use replicator equation (see Schuster and Sigmund (1983); Hofbauer and Sigmund (2003)). We denote by x, y, and z the frequencies of C, D, and R, respectively. Thus, $x, y, z \geq 0$ and $x + y + z = 1$. The replicator equation can be written as

$$\begin{cases} \dot{x} = x(P_C - \bar{P}), \\ \dot{y} = y(P_D - \bar{P}), \\ \dot{z} = z(P_R - \bar{P}), \end{cases} \tag{4}$$

where P_C, P_D, and P_R denote the expected payoffs of C, D, and R, respectively, and $\bar{P} = xP_C + yP_D + zP_R$ gives the average payoff of the entire population. Correspondingly, we have

$$P_i = \sum_{N_C=0}^{N-1} \sum_{N_D=0}^{N-N_C-1} \binom{N-1}{N_C} \binom{N-N_C-1}{N_D} x^{N_C} y^{N_D} z^{N_R} \pi_i,$$

where $i = C, D$, or R.

After taking some calculations, the average payoffs of these three strategies can be, respectively, written as

$$P_C = \frac{rc}{N}(N-1)x + \frac{rc}{N} - c + (N-1)z\alpha,$$
$$P_D = \frac{rc}{N}(N-1)x + pq(N-1)z\alpha - pq(N-1)zh,$$
$$P_R = \frac{rc}{N}(N-1)x - G_R + pq(N-1)yb,$$

where $(N-1)x$ denotes the expected numbers of cooperators among the $N-1$ co-players, and $(N-1)z\alpha$ gives the expected bonus from rewarders.

The population dynamics, including the distribution and stability of equilibrium points, can be studied analytically. For convenience, we introduce the abbreviation $\sigma = \frac{c-rc/N}{(\alpha-pq\alpha+pqh)(N-1)}$, $\xi = \frac{rc/N-c+G_R+\alpha(c-rc/N)/(\alpha-pq\alpha+pqh)}{pq(N-1)b}$, $\varphi = \frac{c-rc/N-G_R}{(N-1)\alpha}$, and $\varsigma = \frac{pq(N-1)b-G_R}{(\alpha-h+b)(N-1)pq}$.

We first investigate the distribution of interior equilibrium points of replicator equation (4). Solving $P_C = P_D$ results in $z = \sigma$. Similarly, by solving $P_C = P_R$, we have $y = \xi$. Thus, there is an interior equilibrium point $(1-\sigma-\xi, \xi, \sigma)$ when $0 < \xi < 1$, $\sigma < 1$, and $0 < \xi + \sigma < 1$.

Then we investigate the dynamics on each edge of the simplex S_3. On the edge CD we have $z = 0$, resulting in $\dot{y} = y(1-y)(P_D - P_C) = y(1-y)(c - \frac{rc}{N}) > 0$. Thus the direction of the dynamics goes from C to D. On the edge CR we have $y = 0$, resulting in $\dot{x} = x(1-x)(P_C - P_R) = x(1-x)[\frac{rc}{N} - c + (N-1)z\alpha + G_R]$. Solving $P_R = P_C$ results in $z = \varphi$. Thus there exists a boundary equilibrium point $(1-\varphi, 0, \varphi)$ when $0 < \varphi < 1$ is satisfied. On the edge DR we have $x = 0$, and the replicator equation changes to $\dot{z} = z(1-z)(P_R - P_D)$. Solving $P_R = P_D$ results in $z = \varsigma$. Thus there exists a boundary equilibrium point $(0, 1-\varsigma, \varsigma)$ when $0 < \varsigma < 1$ is satisfied.

Thus, there are at most six equilibria $(x, y, z) = (1, 0, 0)$, $(0, 1, 0)$, $(0, 0, 1)$, $(1-\varphi, 0, \varphi)$, $(0, 1-\varsigma, \varsigma)$, and $(1-\sigma-\xi, \xi, \sigma)$ in the replicator equation (4). Here the first five are boundary fixed points, while the last one is an interior fixed point. Furthermore, we study the stability of these equilibria by judging the sign of the eigenvalues of Jacobian matrix (see Khalil (1996)). Here we set that

$$\begin{cases} f(x,y) = x[(1-x)(P_C - P_R) - y(P_D - P_R)], \\ g(x,y) = y[(1-y)(P_D - P_R) - x(P_C - P_R)]. \end{cases} \tag{5}$$

Then the Jacobian matrix of the system is

$$J = \begin{bmatrix} \frac{\partial f(x,y)}{\partial x} & \frac{\partial f(x,y)}{\partial y} \\ \frac{\partial g(x,y)}{\partial x} & \frac{\partial g(x,y)}{\partial y} \end{bmatrix}, \tag{6}$$

where

$$\begin{cases} \frac{\partial f(x,y)}{\partial x} = (1-x)[rc/N - c + (N-1)(1-x-y)\alpha \\ + G_R - pq(N-1)yb] - y[pq(N-1)(1-x-y)\alpha \\ - pq(N-1)(1-x-y)h + G_R - pq(N-1)yb] \\ + x\big\{ - [rc/N - c + (N-1)(1-x-y)\alpha + G_R \\ - pq(N-1)yb] + (1-x)[-(N-1)\alpha] - y[-pq(N-1)\alpha \\ + pq(N-1)h]\big\}, \\ \frac{\partial f(x,y)}{\partial y} = x\big\{(1-x)[-(N-1)\alpha - pq(N-1)b] - [pq(N \\ - 1)(1-x-y)\alpha - pq(N-1)(1-x-y)h + G_R - pq(N \\ - 1)yb] - y[-pq(N-1)\alpha + pq(N-1)h - pq(N-1)b]\big\}, \\ \frac{\partial g(x,y)}{\partial x} = y\big\{(1-y)[-pq(N-1)\alpha + pq(N-1)h] \\ - [rc/N - c + (N-1)(1-x-y)\alpha + G_R - pq(N-1)yb] \\ - x[-(N-1)\alpha]\big\}, \\ \frac{\partial g(x,y)}{\partial y} = (1-y)[pq(N-1)(1-x-y)\alpha - pq(N \\ - 1)(1-x-y)h + G_R - pq(N-1)yb] - x[rc/N - c \\ + (N-1)(1-x-y)\alpha + G_R - pq(N-1)yb] \\ + y\big\{ - [pq(N-1)(1-x-y)\alpha - pq(N-1)(1-x-y)h \\ + G_R - pq(N-1)yb] + (1-y)[-pq(N-1)\alpha + pq(N \\ - 1)h - pq(N-1)b] - x[-(N-1)\alpha - pq(N-1)b]\big\}. \end{cases}$$

Accordingly, the concrete expressions of the Jacobian matrix at the above six fixed points can be respectively written as

(1) For $(x, y, z) = (0, 0, 1)$, the Jacobian matrix is

$$J = \begin{bmatrix} a_1 & 0 \\ 0 & pq(N-1)(\alpha - h) + G_R \end{bmatrix},$$

where $a_1 = G_R + \frac{rc}{N} - c + (N-1)\alpha$.

(2) For $(x, y, z) = (1, 0, 0)$, the Jacobian matrix is

$$J = \begin{bmatrix} c - \frac{rc}{N} - G_R & -G_R \\ 0 & c - \frac{rc}{N} \end{bmatrix}.$$

(3) For $(x, y, z) = (0, 1, 0)$, the Jacobian matrix is

$$J = \begin{bmatrix} \frac{rc}{N} - c & 0 \\ a_2 & pq(N-1)b - G_R \end{bmatrix},$$

where $a_2 = pq(N-1)b - G_R + c - \frac{rc}{N}$.

(4) For $(x, y, z) = (0, 1-\varsigma, \varsigma)$, then the Jacobian matrix is

$$J = \begin{bmatrix} a_3 & 0 \\ a_4\ \varsigma(1-\varsigma)pq(N-1)(h-b-\alpha) \end{bmatrix},$$

where $a_3 = G_R + \frac{rc}{N} - c - pq(N-1)b + (N-1)\varsigma(\alpha + pqb)$ and $a_4 = \varsigma(1-\varsigma)(N-1)pq(h-\alpha) - (1-\varsigma)[\frac{rc}{N} - c + (N-1)\varsigma\alpha + G_R - pq(N-1)(1-\varsigma)b]$.

(5) For $(x, y, z) = (1-\varphi, 0, \varphi)$, then the Jacobian matrix is

$$J = \begin{bmatrix} -\varphi(1-\varphi)(N-1)\alpha & a_5 \\ 0 & a_6 \end{bmatrix},$$

where $a_5 = \varphi(1-\varphi)[-(N-1)\alpha - pq(N-1)b] - (1-\varphi)[pq(N-1)\varphi(\alpha - h) + G_R]$ and $a_6 = pq(N-1)\varphi(\alpha - h) + G_R$.

(6) For $(x, y, z) = (1-\sigma-\xi, \xi, \sigma)$, we define the equilibrium point as (x^*, y^*, z^*) hereafter, then the Jacobian matrix is

$$J = \begin{bmatrix} a_7 & a_8 \\ a_9 & a_{10} \end{bmatrix},$$

where

$$\begin{cases} a_7 = -x^*(1-x^*)(N-1)\alpha - x^*y^*(N-1)pq(h-\alpha), \\ a_8 = -x^*(N-1)[(1-x^*)(pqb+\alpha) - y^*pq(\alpha + b - h)], \\ a_9 = y^*(1-y^*)(N-1)pq(h-\alpha) + x^*y^*(N-1)\alpha, \\ a_{10} = -y^*(N-1)[(1-y^*)pq(\alpha + b - h) - x^*(pqb+\alpha)]. \end{cases}$$

We then distinguish six substantially different parameter regions based on the above theoretical analysis, and the corresponding algebra derivation is not presented here because of the limited space. We accordingly summary the distribution and stability of equilibria as follows.

(1) In the condition of $G_R < c - rc/N < (pqh + \alpha - pq\alpha)[pq(N-1)b - G_R]/[pq(\alpha - h + b)]$, the replicator equation (4) has six equilibria which are $(1, 0, 0)$, $(0, 1, 0)$, $(0, 0, 1)$, $(1-\varphi, 0, \varphi)$, $(0, 1-\varsigma, \varsigma)$, and $(1-\sigma-\xi, \xi, \sigma)$, respectively. According to the sign of the largest eigenvalues, the first five are unstable, and the last one is a stable fixed point.

(2) In the condition of $c - rc/N < G_R < (pqh + \alpha - pq\alpha)[pq(N-1)b - G_R]/[pq(\alpha - h + b)]$ or $c - rc/N < (pqh + \alpha - pq\alpha)[pq(N-1)b - G_R]/[pq(\alpha - h + b)] < G_R$, the replicator equation (4) has five equilibria which

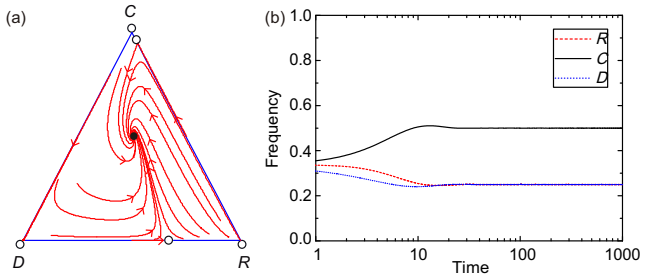

Fig. 1. The system converges to the mixture equilibrium of the three strategists. The triangle represents the state space, $\Delta = \{(x, y, z) : x, y, z \geq 0 \text{ and } x + y + z = 1\}$. Filled circles represent stable fixed points whereas open circles represent unstable fixed points. Panel (a) depicts that a stable interior equilibrium appears in the simplex S_3, which means that these three strategists coexist in the population. Panel (b) depicts time series of the frequencies of three strategies C (pure cooperators, black solid line), R (pure pool rewarders, red dash line), and D (pure defectors, blue dot line). Initial conditions are: $(x, y, z) = (1/3, 1/3, 1/3)$. Parameters: $N = 5$, $r = 3$, $c = 1$, $G_R = 0.3$, $\alpha = 0.5$, $h = 0.4, b = 0.4, q = 1$, and $p = 1$.

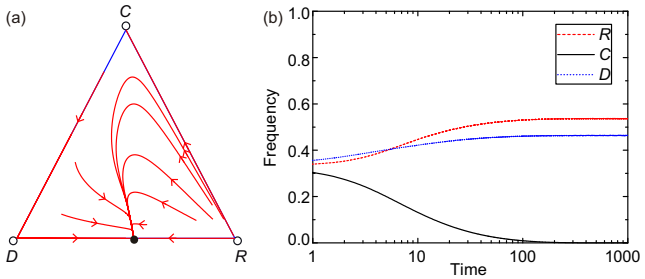

Fig. 2. The system converges to mixture equilibrium of defectors and rewarders. Panel (a) depicts that all the interior curves of simplex S_3 coverage to the fixed point on DR edge, which means that defectors and pure pool rewarders can coexist steadily in the population. Panel (b) depicts time series of the frequencies of three strategies C, D, and R. Initial conditions are: $(x, y, z) = (1/3, 1/3, 1/3)$. Parameters: $N = 5$, $r = 3$, $c = 1$, $G_R = 0.5$, $\alpha = 0.5$, $h = 0.1, b = 0.8, q = 1$, and $p = 0.8$.

are $(1, 0, 0)$, $(0, 1, 0)$, $(0, 0, 1)$, $(0, 1 - \varsigma, \varsigma)$, and $(1 - \sigma - \xi, \xi, \sigma)$, respectively. According to the sign of the largest eigenvalues, the first four are unstable, and the last one is a stable fixed point.

(3) In the condition of $G_R < (pqh + \alpha - pq\alpha)[pq(N-1)b - G_R]/[pq(\alpha - h + b)] < c - rc/N < (N-1)\alpha + G_R$ or $0 < (pqh + \alpha - pq\alpha)[pq(N-1)b - G_R]/[pq(\alpha - h + b)] < G_R < c - rc/N < (N-1)\alpha + G_R$, the replicator equation (4) has five equilibria which are $(1, 0, 0)$, $(0, 1, 0)$, $(0, 0, 1)$, $(1-\varphi, 0, \varphi)$, and $(0, 1-\varsigma, \varsigma)$, respectively. According to the sign of the largest eigenvalues, the first four are unstable, and the last one is a stable fixed point.

(4) In the condition of $0 < (pqh + \alpha - pq\alpha)[pq(N-1)b - G_R]/[pq(\alpha - h + b)] < c - rc/N < G_R$, the replicator equation (4) has four equilibria which are $(1, 0, 0)$, $(0, 1, 0)$, $(0, 0, 1)$, and $(0, 1-\varsigma, \varsigma)$, respectively. According to the sign of the largest eigenvalues, the first three are unstable, and the last one is a stable fixed point.

(5) In the condition of $(pqh + \alpha - pq\alpha)[pq(N-1)b - G_R]/[pq(\alpha - h + b)] < 0 < G_R < c - rc/N < (N-1)\alpha + G_R$, the replicator equation (4) has four equilibria which are $(1, 0, 0)$, $(0, 1, 0)$, $(0, 0, 1)$, and $(1 - \varphi, 0, \varphi)$, respectively. According to the sign of the largest eigenvalues, the equilibria $(1, 0, 0)$, $(0, 0, 1)$, and $(1 - \varphi, 0, \varphi)$ are unstable, while $(0, 1, 0)$ is stable.

(6) In the condition of $(pqh + \alpha - pq\alpha)[pq(N-1)b - G_R]/[pq(\alpha - h + b)] < 0 < c - rc/N < G_R$, the replicator equation (4) has three equilibria which are $(1, 0, 0)$, $(0, 1, 0)$, and $(0, 0, 1)$, respectively. According to the sign of the largest eigenvalues, the equilibria $(1, 0, 0)$ and $(0, 0, 1)$ are unstable, while $(0, 1, 0)$ is stable.

In the following, we provide three representative numerical examples to confirm the above theoretical analysis. The evolutionary dynamics of these three strategies are depicted in the state space $S_3 = \{(x, y, z) : x, y, z \geq 0, x + y + z = 1\}$. Accordingly, the three homogeneous states C $(x = 1)$, D $(y = 1)$, and R $(z = 1)$ correspond to three corners of the simplex S_3.

3.1 The mixture equilibrium of the three strategies

We first present the evolutionary dynamics of the system for $G_R < c - rc/N < (pqh + \alpha - pq\alpha)[pq(N-1)b - G_R]/(pqb + pq\alpha - pqh)$. In this situation, the system has all six equilibrium points (see Fig.1(a)). The interior equilibrium point is stable, and all interior orbits converge to this point, irrespective of the initial conditions. Besides, there is an unstable boundary equilibrium on the edge CR and DR, respectively. A representative time evolution of the frequencies of these three strategies is plotted in Fig.1(b). It suggests that all mentioned three strategists can coexist when the system reaches the steady state. Furthermore, for the given model parameters, the frequency of cooperators is the highest, while the frequencies of rewarders and defectors are both lower in the steady state.

3.2 The mixture equilibrium of defection and pure reward

When $0 < (pqh + \alpha - pq\alpha)[pq(N-1)b - G_R]/(pqb + pq\alpha - pqh) < c - rc/N < G_R$, there is no interior equilibrium point in the state space. As shown in Fig.2(a), there are four equilibrium points in the state space. All interior curves converge to the boundary equilibrium point on DR edge, irrespective of the initial conditions. This result reveals that the system will remain in a mixed state of pure rewarders and defectors where cooperation is extinct. In Fig.2(b), we provide a specific example of how the frequency of three strategies varies over time when the initial fractions of these three strategies are the same. The result shows that rewarders and defectors can coexist steadily, while the frequency of cooperators is reduced to zero.

3.3 The global attractor D

Finally, we present a numerical example to confirm that the equation system can converge to full defection state.

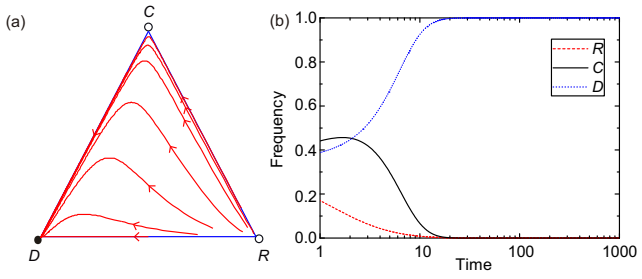

Fig. 3. The system converges to a homogeneous state in which all play D. Panel (a) depicts that the interior curves converge to the vertex D. Panel (b) depicts time series of the frequencies of C, D, and R. Initial conditions are: $(x, y, z) = (1/3, 1/3, 1/3)$. Parameters: $N = 5$, $r = 3$, $c = 1$, $G_R = 0.8$, $\alpha = 0.8$, $h = 0.3, b = 0.3, q = 0.5$, and $p = 1$.

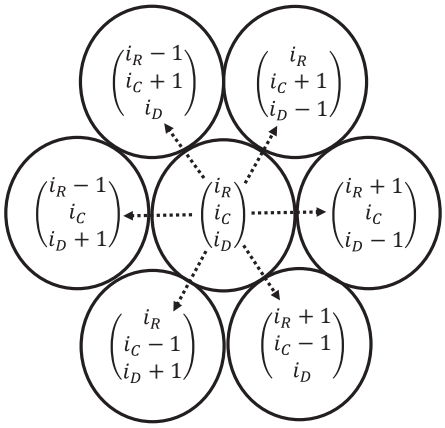

Fig. 4. An example of a local phase space and possible transitions in one-step process. The vector in the central circle represents the configuration of the population at a given time t. The vectors of the surrounding six circles represent the adjacent six configurations that may be reached at the next time step, respectively. The sum of these vectors is the so-called gradient of selection of configuration $\mathbf{i}$ (see Vasconcelos et al. (2013)).

Supposing $(pqh+\alpha-pq\alpha)[pq(N-1)b-G_R]/(pqb+pq\alpha-pqh) < 0 < c - rc/N < G_R$, then the interior fixed point and the boundary equilibrium point on DR and CR edge cannot appear. As shown in Fig.3(a), there are three fixed points in the simplex S_3. All interior orbits converge to the vertex D, which is a global attractor. Besides, no boundary equilibrium point can be detected on the three edges. Figure 3(b) shows that the frequency of D reaches one with time.

4. EVOLUTIONARY DYNAMICS IN FINITE WELL-MIXED POPULATIONS

However, the real population of human society is usually finite and rather small, which is contrary to the hypothesis underlying the dynamics portrayed in Sec. 3. In finite populations, stochastic effects may play an important role in evolutionary dynamics. To characterize the stochastic effects, we assume that there are i_C Cs, i_R Rs, and $Z - i_C - i_R$ in a finite well-mixed population of Z individuals. N players are randomly sampled from the population and assembled into a group to participant in the PGG. It is worth noting that sampling of individuals is no longer binomial, but obeys a hypergeometric distribution. Thus the average payoffs of C, D, and R in a configuration $\mathbf{i} = \{i_R, i_C\}$ using a multivariate hypergeometric sampling can be, respectively, written as

$$f_C(i_R, i_C) = \sum_{N_C=0}^{N-1} \sum_{N_R=0}^{N-N_C-1} \binom{Z-1}{N-1}^{-1} \binom{i_C-1}{N_C} \binom{i_R}{N_R} \binom{Z-i_C-i_R}{N-N_C-N_R-1} \pi_C,$$

$$f_D(i_R, i_C) = \sum_{N_C=0}^{N-1} \sum_{N_R=0}^{N-N_C-1} \binom{Z-1}{N-1}^{-1} \binom{i_C}{N_C} \binom{i_R}{N_R} \binom{Z-i_C-i_R-1}{N-N_C-N_R-1} \pi_D,$$

$$f_R(i_R, i_C) = \sum_{N_C=0}^{N-1} \sum_{N_R=0}^{N-N_C-1} \binom{Z-1}{N-1}^{-1} \binom{i_C}{N_C} \binom{i_R-1}{N_R} \binom{Z-i_C-i_R}{N-N_C-N_R-1} \pi_R,$$

where π_C, π_D, and π_R, respectively, denote payoff values of C, D, and R, which are defined by Eqs.(1-3).

We adopt a stochastic birth-death process combined with the pairwise comparison rule to describe how the number of individuals adopting a given strategy evolves in time for finite populations (see Vasconcelos et al. (2013); Hilbe et al. (2018)). Concretely, we assume that time evolves in discrete steps. At every time step, a randomly selected player L is updated. With probability μ, L undergoes a mutation, and he/she adopts a strategy randomly from the remaining available strategies space. While with probability $1-\mu$, another randomly selected player U acts as a role model for player L. The learner compares his/her payoff (f_L) with that of the role model (f_U), and individual L decides to adopt U's strategy with a probability

$$\frac{1}{1+e^{\beta(f_L-f_U)}}, \tag{7}$$

where $\beta \geq 0$ is the intensity of selection. In the limit case $\beta \to \infty$, the more successful strategy of the role model always succeeds in enforcing his/her strategy to the learner L, but never otherwise. $\beta \to 0$ indicates the so-called weak selection limit where strategy adoption becomes random independently of the payoff differences. In between these extremes, for a finite value of β, it is likely that a better performing strategy is imitated.

Therefore, the probability to choose an individual out of i_{S_k} players with strategy S_k and transform it into a different strategy S_l, under the pairwise comparison rule with an arbitrary mutation rate μ, is given by

$$T_{S_k \to S_l} = (1-\mu)\Big[\frac{i_{S_k}}{Z}\frac{i_{S_l}}{Z-1}(1+e^{\beta(f_{S_k}-f_{S_l})})^{-1}\Big] + \mu\frac{i_{S_k}}{(d-1)Z}, \tag{8}$$

where $S_k, S_l = C, D$, or R, and d represents the number of alternative strategies in the strategy space (here $d = 3$).

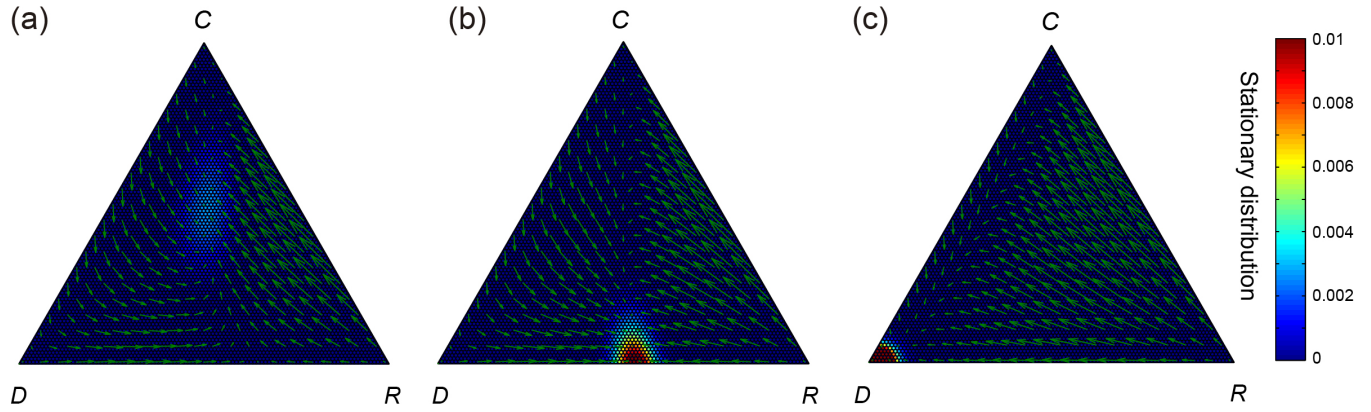

Fig. 5. Dynamics of the system with pure cooperators (C), pure defectors (D), and pure rewarders (R) in the simplex S_3. Panels (a-c) show the stationary distribution and gradient of selection when model parameters are respectively coincide with those in infinite populations (Fig.1, Fig.2, and Fig.3, respectively). Each panel contains all possible configurations of the finite population. Each configuration is represented by a circular dot. Red areas indicate those configurations in which the population spends more time, thus providing a contour representation of the stationary distribution, whose magnitude is shown using the red-green-blue scale indicated. The arrows show the so-called gradient of selection, which provides the most likely direction of evolution from a given configuration. Panel (a) shows that the population spends most of the time in configuration where three strategists coexist. Panel (b) shows that the population spends most of the time in configuration where defectors and pure rewarders coexist. Panel (c) shows that the population spends most of the time in the region of full defection. Parameter values are : $Z = 100, N = 5, r = 3, c = 1, \mu = 1/Z, \beta = 2, G_R = 0.3, \alpha = 0.5, h = 0.4, b = 0.4, q = 1$, and $p = 1$ in (a); $Z = 100, N = 5, r = 3, c = 1, \mu = 1/Z, \beta = 2, G_R = 0.5, \alpha = 0.5, h = 0.1, b = 0.8, q = 1$, and $p = 0.8$ in (b); $Z = 100, N = 5, r = 3, c = 1, \mu = 1/Z, \beta = 2$, $G_R = 0.8, \alpha = 0.8, h = 0.3, b = 0.3, q = 0.5$, and $p = 1$ in (c).

The evolutionary dynamics among C, D, and R can be described by embedded Markov process over a two-dimensional space. The study of Markov process consists in determining its probability density function (PDF) evolution, $p_{\mathbf{i}}(t)$, which provides information on the prevalence of each configuration at time t. Since $\mathbf{i}(t)$ has the Markov property, its transition probability evolves in time according to the discrete time Master Equation (see Van Kampen (1992)). Here, the gain-loss equation allow us to compute $p_{\mathbf{i}}(t)$ as

$$p_{\mathbf{i}}(t+\tau) - p_{\mathbf{i}}(t) = \sum_{\mathbf{i}'} \{T_{\mathbf{i}\mathbf{i}'} p_{\mathbf{i}'}(t) - T_{\mathbf{i}'\mathbf{i}} p_{\mathbf{i}}(t)\}, \qquad (9)$$

where $T_{\mathbf{i}\mathbf{i}'}$ means the transition probability from $\mathbf{i}'$ to adjacent configuration $\mathbf{i}$. Fig.4 illustrates a local representation of the phase space and possible transitions for a two-dimensional one-step process. Similarly, the transitions from neighboring state $\mathbf{i}$ into state $\mathbf{i}'$ are represented by $T_{\mathbf{i}'\mathbf{i}}$. The transition matrix $\Lambda = [T_{\mathbf{i}\mathbf{j}}]^T$ collects the different (transition) probabilities for the population to move from state $\mathbf{j}$ to the other state $\mathbf{i}$. Following previous works (see Imhof et al. (2005); Vasconcelos et al. (2013, 2014)), the so-called stationary distribution is obtained by making the left-hand side of Eq.(9) be equal to zero, which transforms Eq.(9) into an eigenvector search problem, namely, the eigenvector associated with the eigenvalue 1 of the transition matrix Λ.

In addition to the stationary distribution, another important quantity for studying the evolutionary dynamics in finite populations is the gradient of selection. We can compute the gradient of selection which describes the most likely direction of evolution of the system from each configuration $\mathbf{i}$, by resorting to calculating the first coefficient of the Kramers-Moyal expansion of the Master Equation (Details can be found in Helbing (1993); Vasconcelos et al. (2013)). Then the drift vector describing the deterministic part of the dynamics is given by

$$\nabla_{\mathbf{i}} = (T_{\mathbf{i}}^{R+} - T_{\mathbf{i}}^{R-})\mathbf{u}_R + (T_{\mathbf{i}}^{C+} - T_{\mathbf{i}}^{C-})\mathbf{u}_C, \qquad (10)$$

where $\mathbf{u}_R$ and $\mathbf{u}_C$ are two-dimensional unit column vectors with direction of $\frac{\partial \mathbf{i}}{\partial i_R}$ and $\frac{\partial \mathbf{i}}{\partial i_C}$, respectively. Besides, $T_{\mathbf{i}}^{R\pm}$ ($T_{\mathbf{i}}^{C\pm}$) represents the probability to increase (decrease) by one the number of individuals adopting strategy $R(C)$, which can be computed as

$$T_{\mathbf{i}}^{R\pm} = T_{\mathbf{i}\{i_R \pm 1, i_C \mp 1, i_D\}} + T_{\mathbf{i}\{i_R \pm 1, i_C, i_D \mp 1\}}, \qquad (11)$$

$$T_{\mathbf{i}}^{C\pm} = T_{\mathbf{i}\{i_R \mp 1, i_C \pm 1, i_D\}} + T_{\mathbf{i}\{i_R, i_C \pm 1, i_D \mp 1\}}, \qquad (12)$$

where the terms on the right side can be obtained by equation (8). The gradient of selection can be obtained by solving equation (10). The results are shown by the arrows in Fig.5.

Since the stochastic effects can drive the system from the vicinity of one configuration of the state space to the other, which makes an analytical description of the dynamics difficult. Here we provide a numerical investigation to explore the dynamics of the population by calculating the stationary distribution and gradient of selection. In particular, in order to compare with the evolutionary dynamics in infinite populations, we provide three representative numerical examples, where model parameters are consistent with those in Fig.1, Fig.2, and Fig.3, respectively.

As shown in Fig.5(a), defectors can rapidly outcompete cooperators (see arrows), and the population spends most of the time in the interior configurations where three

strategists coexist. This means that cooperation can survive in the population even if the reward system is involved with corruption. Once when we modify the model parameters, the situation is quite different. As shown in Fig.5(b), the population does not spend most of the time in the interior configurations, but instead remains most of the time in the configurations nearby the middle part on DR edge. Finally, for the model parameters of Fig.3, we find that defectors rapidly outcompete cooperators and rewarders, and the population spends most of the time in configurations of pure defectors(see Fig.5(c)).

5. CONCLUSIONS

In summary, we have introduced probabilistic corrupt defectors and rewarders into the PGG and respectively investigated their consequence on the evolution of cooperation and reward in infinite and finite well-mixed populations. We have firstly shown that cooperation can be maintained in an infinite well-mixed population. Both theoretical and numerical results confirm that cooperators, defectors, and rewarders can coexist stably in the population. Furthermore, when the population size is finite, stochastic dynamics show that the population can spend most of the time in the interior region where cooperators can survive. All these results confirm that cooperative behaviors can be maintained even the reward system may suffer from corruption. Besides, it is worth pointing out that in our model, defection will be the most advantageous strategy when corruption is impossible. In this way, the introduction of probabilistic corruption between defectors and rewarders can stabilize cooperation.

Several previous studies have explored the evolution of cooperation and reward in the PGG (see Szolnoki and Perc (2010); Sasaki and Unemi (2011); Szolnoki and Perc (2012)). We have made three key assumptions that differentiate the current work from previous studies. First, we propose a different form of pool reward, in which rewarders need to contribute a constant cost to the reward pool, just like pool punishment (see Sigmund et al. (2010)). Second, we consider that a single individual can afford to perform either cooperation or reward, but not both. Third, we introduce corruption into the PGG, in which defectors can choose to bribe the rewarders probabilistically, and meanwhile rewarders will stochastically receive bribes from defectors. All these differences in model assumptions lead to interesting evolutionary outcome, that is, the coexistence of cooperation, defection, and reward can be globally stable.

REFERENCES

Abdallah, S., Sayed, R., Rahwan, I., LeVeck, B. L., Cebrian, M., Rutherford, A., and Fowler, J. H. (2014). Corruption drives the emergence of civil society. *J. R. Soc. Interface*, 11 : 20131044.

Chen, X., Szolnoki, A., and Perc, M. (2014). Probabilistic sharing solves the problem of costly punishment. *New J. Phys.*, 16 : 083016.

Chen, X., Szolnoki, A., and Perc, M. (2015). Competition and cooperation among different punishing strategies in the spatial public goods game. *Phys. Rev. E*, 92 : 012819.

De Quervain, D. J., Fischbacher, U., Treyer, V., and Schellhammer, M. (2004). The neural basis of altruistic punishment. *Science*, 305 : 1254–1258.

Fehr, E. and Gächter, S. (2002). Altruistic punishment in humans. *Nature*, 415 : 137–140.

Gao, L., Wang, Z., Pansini, R., Li, Y. T., and Wang, R. W. (2015). Collective punishment is more effective than collective reward for promoting cooperation. *Sci. Rep.*, 5 : 17752.

Govaert, A., Ramazi, P., and Cao, M. (2017). Convergence of imitation dynamics for public goods games on networks. *2017 IEEE 56th Annual Conference on Decision and Control (CDC)*, pages 4982–4987.

Helbing, D. (1993). Boltzmann-like and Boltzmann-Fokker-Planck equations as a foundation of behavioral models. *Physica A*, 196 : 546–573.

Hilbe, C., Šimsa, Š., Chatterjee, K., and Nowak, M. A. (2018). Evolution of cooperation in stochastic games. *Nature*, 559 : 246-249.

Hofbauer, J. and Sigmund, K. (2003). Evolutionary game dynamics. *Bull. Am. Math. Soc.*, 40 : 479–519.

Hauert, C., Traulsen, A., Brandt, H., Nowak, M. A., and Sigmund, K. (2007). Via freedom to coercion: the emergence of costly punishment. *Science*, 316 : 1905–1907.

Huang, F. and Chen, X., and Wang, L. (2018). Evolution of cooperation in a hierarchical society with corruption control. *J. Theor. Biol.*, 449 : 60–72.

Imhof, L. A., Fudenberg, D., and Nowak, M. A. (2005). Evolutionary cycles of cooperation and defection. *Proc. Natl. Acad. Sci. USA.*, 102 : 10797–10800.

Khalil, H. (1996). *Nonlinear Systems*. Prentice-Hall NJ.

Kawano, Y., Gong, L., Anderson, B. D., and Cao, M. (2017). Evolutionary dynamics of two communities under environmental feedback. *IEEE Contr. Syst. Lett.*, 3 : 254–259.

Liu, L., Chen, X., and Szolnoki, A. (2019). Evolutionary dynamics of cooperation in a population with probabilistic corrupt enforcers and violators. *Math. Models Methods Appl. Sci.*, 29 : 2127–2149.

Liu, L., Chen, X., and Perc, M. (2019). Evolutionary dynamics of cooperation in the public goods game with pool exclusion strategies. *Nonlinear Dynam.*, 97 : 749–766.

Muthukrishna, M., Francois, P., Pourahmadi, S., and Henrich, J. (2017). Corrupting cooperation and how anti-corruption strategies may backfire. *Nat. Hum. Behav.*, 1 : 0138.

Nowak, M. A., Sasaki, A., Taylor, C., and Fudenberg, D. (2004). Emergence of cooperation and evolutionary stability in finite populations. *Nature*, 428 : 646–650.

Nowak, M. A. (2006). Five rules for the evolution of cooperation. *Science*, 314 : 1560–1563.

Perc, M., Jordan, J. J., Rand, D. G., Wang, Z., Boccaletti, S., and Szolnoki, A. (2017). Statistical physics of human cooperation. *Phys. Rep.*, 687 : 1–51.

Rand, D. G., Dreber, A.,Ellingsen, T., Fudenberg, D., and Nowak, M. A. (2007). Positive interactions promote public cooperation. *Science*, 325 : 1272–1275.

Schuster, P. and Sigmund, K. (1983). Replicator dynamics. *J. Theor. Biol.*, 100 : 533–538.

Sigmund, K. and Hauert, C. and Nowak, M. A. (2001). Reward and punishment. *Proc. Natl. Acad. Sci. USA.*,

98 : 10757–10762.

Szolnoki, A. and Perc, M. (2010). Reward and cooperation in the spatial public goods game. *EPL*, 92 : 38003.

Sigmund, K., De Silva, H., Traulsen, A., and Hauert, C. (2010). Social learning promotes institutions for governing the commons. *Nature*, 466 : 861–863.

Sasaki, T. and Unemi, T. (2011). Replicator dynamics in public goods games with reward funds. *J. Theor. Biol.*, 287 : 109–114.

Szolnoki, A. and Perc, M. (2012). Evolutionary advantages of adaptive rewarding. *New J. Phys.*, 14 : 093016.

Van Kampen, N.G. (1992). *Stochastic Processes in Physics and Chemistry*. Elsevier.

Vasconcelos, V. V., Santos, F. C., and Pacheco, J. M. (2013). A bottom-up institutional approach to cooperative governance of risky commons. *Nat. Clim. Change*, 3 : 797–801.

Vasconcelos, V. V., Santos, F. C., Pacheco, J. M., and Levin, S. A. (2014). Climate policies under wealth inequality. *Proc. Natl. Acad. Sci. USA.*, 111 : 2212–2216.

Wang, Q., He, N., and Chen, X. (2018). Replicator dynamics for public goods game with resource allocation in large populations. *Appl. Math. Comput.*, 328 : 162-170.

Wang, S., Chen, X., and Szolnoki, A. (2019). Exploring optimal institutional incentives for public cooperation. *Commun. Nonlinear Sci. Numer. Simul.*, 79 : 104914.

Xiao, Z., Chen, X., and Szolnoki, A. (2020). Leaving bads provides better outcome than approaching goods in a social dilemma. *New J. Phys.*, 22 : 023012.